

Aberystwyth University

The Logic of Biological Inspired Robotics

Hülse, Martin; Lee, Mark

Published in:

Proceedings of the International Symposium on AI-Inspired Biology

Publication date:

2010

Citation for published version (APA):

Hülse, M., & Lee, M. (2010). The Logic of Biological Inspired Robotics. In J. Chappell, S. Thorpe, N. Hawes, & A. Sloman (Eds.), *Proceedings of the International Symposium on AI-Inspired Biology* (pp. 43 - 50)
<http://hdl.handle.net/2160/5697>

General rights

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

The Logic of Robotics Inspired Biology

Martin Hülse and Mark Lee¹

Abstract. Biologically inspired robotics is a well known approach for the design of autonomous intelligent robot systems. Very often it is assumed that biologically inspired models successfully implemented on robots offer new scientific knowledge for biology too. In other words, robots experiments serving as a replacement for the biological system under investigation are assumed to provide new scientific knowledge for biology. This article is a critical investigation of this assumption. We begin by clarifying what we mean by “new scientific knowledge.” Following Karl Popper’s work the *The Logic of Scientific Discovery* we conclude that in general robotic experiments serving as replacement for biological systems can never directly deliver any new scientific knowledge for biology. We further argue that there is no formal guideline which defines the level of “biological plausibility” for biologically inspired robot implementations. Therefore, there is no reason to prefer some kind of robotic setup before others. Any claimed relevance for biology, however, is only justified if results from robotic experiments are translated back into new models and hypotheses amenable to experimental tests within the domain of biology. This translation “back” into biology is very often missing and we will discuss popular robotics frameworks in the context of Brain Research, Cognitive Science and Developmental Robotics in order to highlight this issue. Nonetheless, such frameworks are valuable and important, like pure mathematics, because they might lead to new formalisms and methods which in future might be essential for gaining new scientific knowledge if applied in biology. No one can tell, if and which of the current robotics frameworks will provide these new scientific tools. What we can already say—the main message of this article—is that robot systems serving as a replacement for biological systems won’t be sufficient for the test of biological models, i.e. gaining new scientific knowledge in biology.

1 Introduction

Nature has been an inexhaustible and illuminating source of inspiration for scientists and engineers through all disciplines and it will continue to be so without any doubt. When engineering reaches its limits, nature might have good ideas how to proceed. In the area of autonomous and intelligent robot systems living beings outperform current engineering by far. Therefore not surprisingly a whole body of work in this field is entitled “biologically inspired robotics” where research is focused on copying biological systems. The hope is that this might lead to novel technologies which will help to close the current gap between natural and artificial intelligence. Whether or not this kind of biological inspiration has truly led to better robot systems shall not be the scope of this paper. We rather ask whether an implementation of a biologically inspired model on a robot sys-

tem can have any scientific value for biology. In other words, *can biologists gain new scientific knowledge from biological models implemented on robots?* This is the guiding question of this article.

The answer can have interesting implications. If the answer is positive then robotics or computer science in general could be seen as a discipline which provides scientific knowledge about biological phenomena. Hence, computer science would be part of the empirical sciences, which is in fact an “old dream” of Artificial Intelligence [24].

Once the far reaching consequences of this question has been acknowledged we must carefully proceed in our search for an answer. The objective of this article is to provide some solid ground which might also be a starting point for future and more comprehensive studies of this issue. We begin by explicitly defining what we generally mean by “gaining new scientific knowledge.” As the reader will see, the Philosophy of Science and Karl Popper’s work *The Logic of Scientific Discovery* [18] in particular will provide a coherent definition (Section 2). Applying this definition in the context of biology and robotics, we will come to the conclusion that robotic experiments can never directly deliver any scientific knowledge for biology (Section 3). However, there are frameworks claiming to be of “biological relevance.” We will critically discuss this issue in Section 3 where outline the difference between “new insights” and “new scientific knowledge”. We argue that the relevance of robotic systems for new scientific knowledge in biology can only lie in the potential future impact on biology by developing new formalisms, paradigms and scientific tools leading to advanced studies of biological phenomena. This is explained in Section 4 by concrete examples from Brain Research and Cognitive Science followed by a case study outlining our work in Developmental Robotics (Section 5).

2 Karl Popper’s Logic of Scientific Discovery

Since ancient times philosophers have been puzzled by the problem of how human beings are able to gain knowledge about the world if there is no logical justification that observations made in the past will occur in the future. Therefore, any theory or general law derived from past observations (inductive reasoning) is logically not valid. As a consequence, the truth of a theory inferred by inductive reasoning can not be proven, thus, the truth of such a theory is not ensured. This is called the problem of induction [22].

The problem of induction is addressed by Karl Popper in his main work *The Logic of Scientific Discovery* [18] where he rejects inductive reasoning as a method for theory-building. According to Karl Popper a logically valid foundation for knowledge discoveries in empirical sciences can only be provided by the process of falsification. The statements or hypotheses derived from a theory via deduction can be tested empirically by conducting experiments. If the predictions inferred from a hypothesis do not match with the outcome of

¹ Intelligent Robotics Group, Department of Computer Science, Aberystwyth University, SY23 3DB, Wales, UK, email: {msh, mhl}@aber.ac.uk

the experiment, then the theory must be refuted because it obviously leads to a wrong statement. And a theory which allows the deduction of wrong statements cannot provide a logically valid description of the phenomena under investigation. On the other hand, if the experimental outcome matches with the predictions of the theory then the theory is validated. Nonetheless, a validation does not tell us anything about the outcome of future experiments. Therefore, the truth of the theory remains questionable no matter how many times it has been validated, and moreover, validations not even make them more probable.

Karl Popper points out that scientific knowledge is gained through falsification only. Falsification is the only logically valid way that drives the development of new theories because falsification usually does not lead to the refutation of the theory as a whole. It rather guides a careful revision of it which will lead to a better theory. According to Popper, scientists therefore shall be aiming for experiments that falsify their theories rather than just validate them.

Notice, the origins of a theory, i.e. how human beings develop axioms or formal expression of general laws, is not the subject of Karl Popper's analysis. He even understands this issue as not relevant for the logical foundations of empirical sciences.

Without going into much detail, we have to clarify some terminology first before we can go on discussing Popper's view in the context of robotics and biology. Modern logic defines a theory as a set of sentences of a formal language. A model is an interpretation of the theory that makes all sentences of the theory true.

In empirical sciences models are understood to be interpretations of general laws. These laws are expressed in a form of calculus, like the equations of Newton's three laws of motion [10]. Taking the example from classical physics given in [10] one can say, general laws "are applied to a particular system, e.g. a pendulum, by choosing a special force function, making assumptions about the mass distribution of the pendulum etc. The resulting model then is an interpretation (or realization) of the general law."

Nowadays, the majority of discoveries in empirical sciences are presented as models. As long as a model is derived from an interpretation of a theory (a set of sentences of a formal language) then this model allows the deduction of hypotheses. Thus, models can be subject of a formal falsification process as Popper proposes it.

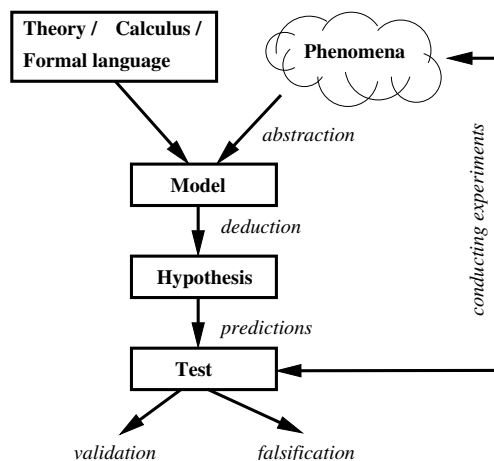


Figure 1. Schema of model-building and testing after Karl Popper [18]

Figure 1 presents a schema describing the process of testing models after Popper [18]. Interested in providing an explanation of some phenomena in a specific domain, scientists derive a model which is grounded in some formal framework. Every model comes with an abstraction of the target phenomena [20]. The level of abstraction is determined by the scientific question as well as by the formal framework (mathematical theory or axiomatic system). A hypothesis can be deduced from the model, which is a logically valid procedure. The derived hypothesis will provide predictions that might be amenable to experiments. If predictions of the model are confirmed by the experiments the model is verified. If not, the model is falsified and must be refuted as an explanation of the targeted phenomena. The latter outcome will lead to a revised and hopefully better model, in other words scientific knowledge will be gained by falsification.

It is important to note that in empirical sciences the *experimental tests of a hypothesis are conducted in the same domain as the phenomena under investigations*. Hypotheses about a biological system are tested by experiments with this biological system. Models of brain phenomena must be tested against data measured from real brains. If experiments are conducted not in the original domain of the target phenomena, then we cannot refer to such experiments as a test of the model for the original phenomena.

3 Robotics in Biology

3.1 The problem of using robots for testing biological models

Regarding the usage of robot systems in biology we understand a robot system as an experimental device where the essential parts are assumed to be a realization or interpretation of a general law or a formal model. The aim of this realization is the experimental tests of a biological model. We assume models as a formal description which allows the deduction of hypotheses. Therefore, the robot system itself is not a model as often proclaimed. It is, including the computational processes running on it, part of an experimental setup. Similar to the example of the pendulum above, a robot is a device to generate data that helps to test the model it instantiates.

In addition we now stress the issue of the experimental domain. Recognising a robot as an experimental device, we have to ask in which domain the experimental data are generated. In other words, where does the data come from? It must be admitted that robot experiments serving as replacement of the biological system under investigation do not generate data in original domain of the biological phenomena. Hence, testing a hypothesis derived from a biological model, i.e. targeting biological phenomena, by exclusively using a robot system is not a logically valid test in the sense of Popper's Logic of Scientific Discovery. Experimenting with a robotic system introduces a new domain, that in general is totally different to the original biological domain of the target phenomena. Thus, any data derived from such robot experiments do not allow any conclusion about the explanatory power of the model with respect to the biological phenomena. In summary, *robotic systems do not provide a logically valid test of a hypothesis about biological systems because the original biological domain is lost*. In consequence, from the point of gaining logically valid scientific knowledge in biology, the use of robot systems cannot be sufficient.

Notice, we do not argue that the change of domain is due to the use of technical equipment. Nowadays, probably any experiment in empirical sciences is in need of sophisticated technical apparatus in order to create controlled conditions and guarantee good measurements. Nevertheless, such devices still generate the requested data

out of the interaction with the target phenomena. This is not the case for a robot system that replaces the biological system. All the data generated by such a robot replacement are totally decoupled and isolated from the domain of the phenomenon a biological model is intended to address. This is what we call the loss of the original domain. Fig. 2 shows an attempt to illustrate this issue.

Leaving the domain B of biological phenomena when testing a hypothesis B derived from the original model has serious consequences. Applying Popper's framework, the predictions of hypothesis B turn into the target phenomena in a robotic domain R. This domain is grounded in a different formal theory which is determined by the computational paradigm, the technology and specific requirements. As a result a new model will be developed which might provide a similar level of abstraction and similar predictions but the test, the experiment, is done in a pure robotic domain. Hence, the loss of the original domain does not provide any conclusions about the explanatory power of model B or R for the phenomena in domain B.

3.2 Why robots at all?

The problem of losing the original domain when performing experiments with robots obviously questions the role of robot experiments in the context of biology. The usage of autonomous robots as scientific method in biology has been an object of debate since their very first appearance in the literature of brain research. Most famous is Grey Walter's book *The living brain* [27], where he introduces robots performing tropisms and even learning tasks. More recently the topic has been raised in articles by Barbara Webb [28, 29]. She proposes seven categories (called dimensions) that are intended to provide a guideline for the evaluation of the relevance of "robot models" for biological phenomena. In her 2009 article [29] she goes even further arguing that the only proof of biological relevance of "robot models" is the proposal of a new hypothesis that can be tested with the real biological systems. The latter statement is much in line with the conclusions in this work. However, in contrast to our argumentation she is not addressing the role of validation and falsification. This lack, it seems to us, might muddy the general demarcation between the

robotic domain and the world of biological phenomena. We deny any logically valid relations between biology and robot experiments, therefore, in general no robot experiment can provide logical valid falsification or validation of a model aimed at biological phenomena. In this respect our claims are much more strict because we say, independent of Webb's seven dimensions (Relevance, Level, Generality, Abstraction, Structural accuracy, Performance match, Medium [28]) models, hypotheses and experiments in the robot domain do not provide any new scientific knowledge for biology. Consequently, Webb's seven dimensions fail in their attempt to provide guidelines for what "good robot models" for biological phenomena are supposed to be.

Only the translation from robotics back into biological experiments allows validation or falsification, and thus, logical valid scientific knowledge. Thus any robot experiment can become a source of formal or informal inspiration that might be translated into the biological domain where it can be tested. Hence, what seems to be quite strict or limiting, i.e. no logical relation between robot domain and biology, provides in fact a much more relaxed view about the tools and methods used to "get inspiration" for good biological models. Since there is no logical relation between the domain of biology and robots, there is no need to introduce limitations in the robotic domain for the sake of its legitimization for biological phenomena. Whether it be a physical robot or a simulation, a detailed high-dimensional model of a specific brain structure or a very abstract and low-dimensional implementation of a dynamical system, they all have their value, if, in the short or the long run, models will be proposed which can be experimentally tested with the biological system under investigation. It is this point where our conclusion from Popper's Logic of Scientific Discovery opposes Barbara Webb's stand for the seven dimensions [28]. We argue, according to the *The Logic of Scientific Discovery* Webb's seven dimension cannot determine the biological relevance of a robot system.

In the following we briefly discuss common strategies used to justify or defend the biological relevance of specific robotic frameworks. Similar issues and many more are discussed in [29], but the reader will see that our proposal lets them appear in a different light.

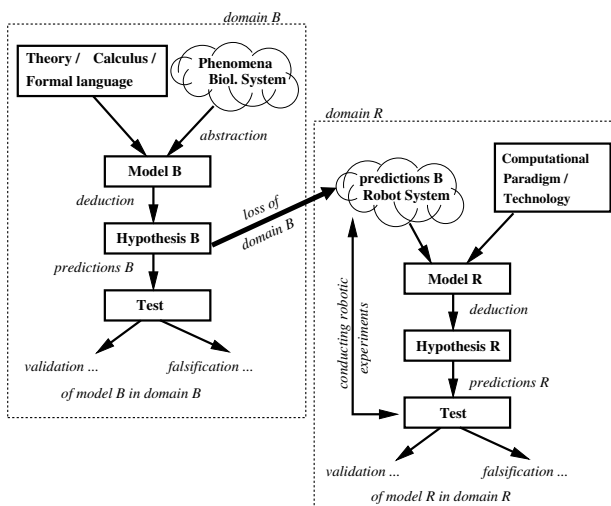


Figure 2. The loss of domain.

3.2.1 Biologically inspired only

Biologically inspired robotics refers to robotic experiments where specific computational paradigms, novel algorithms or optimization processes copy mechanisms observed in biological systems. Sometimes it is claimed that the scientific value of the work is given in terms of a validation of a biological model [3].

According to Popper's framework, however, a validation does not provide any new scientific knowledge. Consequently and without even considering the loss of the biological domain, a successful robot implementation of a biological model does not provide any logically valid statement that this model is to be preferred to any other model targeting the same phenomena.

When presenting biologically inspired robot implementations only, in the best case this can only serve engineering by promoting more efficient methods for new tasks, domains or application; a scientific value for biology is not given.

An interesting question occurs when assuming a biological model confirming a hypothesis which is not confirmed when implemented as a robot model. Apart from the fact that we are not aware of any published work of this kind, the question needs to be addressed. Can biologically inspired robotics experiments serve as a valid falsification of the model for the biological phenomena? We argue that they

cannot. Assume a hypothesis derived from a biological model which is tested in the biological domain as well as it is tested in a setup where a robot replaces the biological system. Given that in the first case the hypothesis is validated while it is falsified in the robotic experiment, then the robotic experiment would hardly be accepted as a falsification of the model. The reason is obviously that a robotic implementation raises a lot of concerns about being a valid representation of the biological system. In other words, the loss of the domain does not allow robotics to be test bed for the falsification of models intended to address biological phenomena.

3.2.2 *Proof of concept*

A proof of concept, also called an existence proof, is usually introduced in order to overcome paradigms that are believed to be the only valid explanation of certain phenomena. As an example the general law: “There is no I without S” can be countered by showing: “There is an I without S”. The most famous example in Artificial Intelligence for this kind of counter example is Rodney Brooks’ article *Intelligence without representation* [5], where he targeted the dominance of the symbolic information processing hypothesis by demonstrating that advanced robot behaviours can be generated without the application of the symbolic information processing paradigm. However, Brooks never said: “All intelligence without symbolic representation.”

With respect to the relevance of existence proofs for an understanding of biological systems we have to recognise that a robotic experiment introduced as an existence proof in a specific robotic domain does not provide any more than the implementation of a biologically inspired model (see discussion above). Hence, in the best case an existence proof serves engineering but not empirical sciences.

3.2.3 *Demonstrations of principles*

W. G. Walter’s robot system were already mentioned. Interestingly, these robot systems were promoted as “part of a mirror of the brain” [27] because they demonstrate principles that must be considered when investigating real brains. In Cybernetics such a principle is the “feedback loop in which the environment is a component.” In this context, robots or machines are seen as a valid and illuminating “mimicry of life”.

Obviously, Walter presented his robots as devices for the validation of a biological model. As we argue this is questionable since the biological domain is lost. However, the issue of “demonstration” includes an additional quality. Demonstrating a principle in a very abstract form gives evidence about the involvement of mechanisms that might be an essential part of many phenomena in different domains. These mechanisms might therefore be worth studying in a formal way, i.e. mathematical framework for their own sake, without the need to refer back to the biological target phenomena.

Whether or not the results of these studies are in the short or the long term helpful to understand specific phenomena of biological systems, this can only be answered by “explicitly translating” the results back into models, hypotheses and experimental tests in the biological domain. Walter (and others) often pursued such model translation exercises as they were primarily interested in neuroscience.

3.2.4 *Support the development of new theoretical frameworks, paradigms, and tools*

The aspect of robots as demonstrators of principles is closely coupled with the development of new theoretical frameworks. Studying

principles, like feedback loops, in a purely formal framework bears the danger that formalisms are evolved which lose any match with real systems. Using robot systems can serve as a first kind of reality check. Demonstrating the outcome of new formalisms on a robot system gives *insights* about the constraints and valid boundary conditions of a formalism. In this respect we understand Webb’s statement that robot experiments provide insights “by working on real problems in real environment” [28]. Notice, we understand *new insights* as different from *new scientific knowledge*. Only if insights gained from robotics experiments are translated back into biological models and tested in the biological domain, they can become new scientific knowledge in biology.

3.2.5 *Robots as tools for the development of new hypotheses amenable to experimental test*

The best way to motivate a new formalism or theory for the old unsolved problems is the ability to infer new models which allow the deduction of already known and new predictions that can be tested experimentally. Real robots, simulations or even a set of equations are promoted as tools that help us to provide new testable hypotheses. Indeed, robots interacting within unconstrained environments or computer simulations often produce unforeseen but reproducible results. These unexpected experimental outcomes might be seen as an inevitable consequences of the model applied and therefore are claimed to be new predictions or hypotheses. Nevertheless, only the experimental test of these predictions in the biological domain allows conclusions whether or not “the new hypothesis” is just a superficial effect of the non-biological domain.

Moreover, the test of a new hypothesis for biological systems derived from experiments in non-biological domains involves an additional problem. If the “new hypothesis” cannot be derived from the biological model in a logically valid way—e.g. without computer simulations or experiments in the robotics domain—then the “new hypothesis” is not a logically valid deduction. Consequently, an experimental test of such a hypothesis makes no sense at all because one does not know whether this hypothesis is a logical consequence of the biological model.

4 **Current robotic frameworks in brain research and cognitive science**

In the following we examine popular robotic frameworks in brain research and cognitive science where we briefly discuss how they match with our five “categories of justification of biological relevance.”

4.1 **Brain Based Devices**

Following specific design principles a brain-based device, introduced and promoted by G. M. Edelman [9], is a physical robot system interacting autonomously with its environment. The control structure of the system is provided by “simulated nervous systems”, in fact large-scale artificial neural networks. In order to allow comparison with real brain-data the design of the artificial neural networks tries to “reflect the brain’s anatomy and physiology” [14]. A series of robot experiments demonstrating a variety of adaptation tasks driven by neural group selection principles is for instance given in [14] where it is stated that brain-based devices serve as “a heuristic for testing theories of brain function.” With respect to “testing” it must be admitted that although the hypotheses and predictions are formulated

in the terminology of brain science they are experimental tested with robot systems only. In this sense the experiments presented have primarily the character of a validation of a robotic model.

Regarding the validation of a biological model it is claimed that the results of the robot experiments are in accordance with brain data from literature. These data, however, were published over ten years earlier. Thus, it is hard to see how the introduced robotic experiments have led to a new hypothesis.

Although strictly designed in relation to a specific level of abstraction in brain science, the experiments of the brain-based devices remain in the robotics domain. Moreover, experiments have the character of validation. It is for instance shown that the “binding problem” can be solved in a particular robot-environment system by synchronisation mechanisms. But as long as these results are not transferred into models and new testable hypotheses for real brains, the scientific value for brain research remains questionable.

From an engineering perspective brain-based devices might offer alternatives to symbolic information processing strategies for the development of advanced robot systems. As the authors emphasise brain-based devices “provide groundwork for the development of intelligent machines that follow neurobiological rather than computational principles in their construction” [14].

In conclusion brain-based devices can be seen as tools for the validation of biologically inspired models of neural information processing and the demonstration of principles (like synchronisation). An explicit formulation of biological models leading to new hypotheses testable in real brains has not been shown yet.

4.2 Conceptual neural evolution

Understanding the complexity of the human brain by investigating how this complexity is rooted in “simpler complexities” of “our evolutionary cousins” is proposed by M. A. Arbib as the framework of conceptual neural evolution. This approach emphasises the principle of “evolutionary refinement” where new behavioural competences emerge from the modulation of already given but simpler behaviours [1]. Arbib states explicitly that the models developed in this framework are derived from and tested with data from real animals. Hence, biologically inspired robot systems are a side effect of this research which “will advance the design principles for a new generation of biologically inspired robots” [1]. One example, a computational model for path planning [2], is explicitly mentioned as a source of inspiration for behaviour-based robotics [1]. Hence, within conceptual neural evolution Arbib sees robot systems as demonstrators of principles of “evolutionary refinement” only. As a source of inspiration for the development of brain models robotics plays no part.

4.3 Evolving Minimal Cognitive Agents

Within the context of cognitive science the application of evolutionary algorithms to generate robot behaviour for real or simulated robot systems is motivated as the evolution of “minimally cognitive behaviour” [4]. Such experiments are meant to be a step towards abstracting general principles of cognition that can be explicitly expressed within the formalism of Dynamical Systems Theory [26]. Since the process of artificial evolution allows the generation of low-dimensional neural circuits, the resulting control structures can be analytically investigated in a reasonable depth.

Scientists in this area, like Randal Beer, admit that the evolution of minimally cognitive behaviour is a result of an abstraction process

which allows no comparison with or empirical predictions for biological phenomena. Thus, the main contribution of this framework is to provide concrete examples within the specific formalisms of Dynamical Systems Theory, because “only by working through the details of tractable concrete examples of the cognitive phenomena of interest that we can begin to glimpse whatever general theoretical principles may exist and whatever mathematical and computational tools will be necessary to formulate them” [4].

Consequently, this approach fits into the category which we have introduced above as the “support of the development of new theoretical frameworks, paradigms and tools.” Once it is explicitly stated that the “evolution of minimally cognitive behaviour” is totally decoupled from the biological phenomena, we are comfortable with following Beer’s argumentation about the possible biological relevance of such studies in the near or far future.

5 A case study from Developmental Robotics

Among psychologists there is a longstanding controversy regarding the origins of many brain functions and observed behaviours. On one hand, the nativists argue that much basic function is innate and is thus specified genetically and becomes hardwired during neural maturation. On the other side, the empiricists claim that much apparently innate competence is instead formed through experience or interaction with the environment and hence involves some kind of adaptation or learning. Much of this debate centres on infant development, particularly the early growth of sensory-motor skills and cognition. Some behaviours, such as the early reflexes, are clearly innate because they have no opportunity for learning and are immediately available. Other functions, such as smiling and responding to faces, are less easily analysed and there are arguments on both sides for their emergence and growth. We now discuss a model of eye movement that illustrates how robotic models may shed some light on the general principles involved in one small aspect of this general problem. It is important to note that our model is not able to comment on the biological structure and validate or falsify particular features, but it is possible to expose some principles of logical necessity that must be present (or absent) in both the robot model and the biological system.

The human newborn is helpless at birth but has available a range of automatic sensory-motor reflex behaviours and also some primitive but potentially fundamental competences that are precursors for cognitive growth. These include basic head orientation to audio stimuli [16] and saccading with the eyes to look at a visual stimulus [21]. Because newborns are able to saccade immediately (they do not have smooth pursuit eye control) this has often been taken as evidence that this is an innate function. We have built a robot system that very quickly learns how to saccade and this demonstrates the possibility that the fundamental process of visual saccading to a peripheral stimulus could be learned rather than innate. If such rapid learning does occur it would be quite difficult to detect but it could be very significant evidence for the empiricist stance [25].

In computer vision or robotics, if a camera (robot “eye”) is to be moved in order to bring an object on the periphery of the image towards the centre then this is usually achieved by complex calculations involving knowledge of the geometry of the camera location, the scene structure and the parameters of the optical setup [30]. Even in systems that try to minimise the engineering of this task there is the need to calibrate the geometric and optical configuration [8].

As part of our work in Developmental Robotics we have built a model that does not need any prior information and is self-

calibrating. Through a process of learning the model essentially performs a kind of continuous calibration in which the mapping between image movement and eye movement is learned. The details are described elsewhere [7] but the method is straightforward. Our robot eye is implemented as two sensory functions: image processing software is used to simulate a periphery sensor and a centre or foveal sensor. The periphery sensor detects the position of any phasic changes in the visual periphery area. The centre sensor analyses any objects (i.e. colour blobs) in the foveal region of the visual field. This configuration is similar to the human vision system which drives saccades towards peripheral stimuli in order to fixate the foveal processing machinery onto targets that attract attention. The main control issue for this ocular-motor system is a sensory-motor coordination problem: what are the necessary motor variables to drive the eye so that a specific sensed region in the image is moved to the centre of the image? This means a transform must be known for converting image space displacement vectors into eye gaze space movement vectors. We designed a simple algorithm that can learn this transform as a mapping between the spaces. Local spatial correspondences are established on the basis of experience and are entered into the mapping in a continuous, cumulative and incremental process. The underlying assumptions are that (a) the mapping may not be a simple or linear relationship and (b) the position of the eyeball can not be related to image data until after birth (as vision is ineffective in the womb [23]) and therefore learning should start from zero prior knowledge.

If we start with no prior information then when the first peripheral target appears there will be no appropriate motor values known and some spontaneous movement must be made in order to gain some experience. When eventually the eye finds the foveal region then the parameters can be recorded in the mapping (initial peripheral location and final angle of gaze) for future reference. After a while sufficient regions become mapped so that only a few movements are needed before finding the target, and eventually the map becomes fully populated so that any target can be brought to the centre in a one step saccade. We observed three distinct qualitatively different behaviours emerge during this learning process: at first random Brownian-like wandering appeared, followed by more directed multi-step paths that appear to involve corrections, and finally single accurate saccades straight to the centre. These behavioural patterns emerged successively although not in completely strict sequence.

The most significant results from this experiment are: the system learns very rapidly; it does not require any calibration process or prior knowledge; it continuously adapts to correct errors; and displays qualitatively different stages in its behaviour which emerge during learning. Perhaps the most important result is the extremely fast learning speed – which was much faster than current neural network based approaches. The simplicity of the method is important in this regard, as the requirement for fast performance, in both saccading and learning, rules out complex computations and the speed of neural processes limits the number of steps that can be involved [17]. We found that some 200 saccade attempts were sufficient to build a near complete mapping, and if this is converted into infant activity the figures show that a saccade map might be built in a matter of hours, or at the most a few days.

Other results from our model show compatibility with the known data on early infant eye activity. Specific infant data on the size, eccentricity and average number of saccades, corroborate the behaviours observed [21, 11]. The accuracy of saccades with our model is a good match with reported infant accuracy [15, 13]. The average error in saccading to a given image location is 0.35 of the tolerance spacing between points in the mapping.

To return to the role of such models, we now see that it is *logically possible* that newborn saccadic eye behaviour could be learned rather than innate. This learning could be so rapid that it might not be easy to detect and the onus is now on psychology and neuroscience to investigate for new evidence. Others have reported similar fast learning in infants across a range of activities [19] and it has been suggested that the preference of newborns to orient towards faces could be learned very rapidly, even in the first six minutes of life [6]. Face orientation is another behaviour that is generally believed to be innate.

With this new viewpoint we can now reflect that it is quite difficult to see how accurate saccades could be created pre-wired before birth. The concrete and specific information needed to coordinate eyeball movement with movement on the image (i.e. movement *of* the image with movement *within* the image) depends upon the physical and optical properties of each particular eye and individual differences in anatomy and geometry. While the proprioceptive feedback from eyeball muscles can be calibrated in the womb, there are no detailed visual signals available prenatally for similar calibration of the image space [23]. And so it seems inconceivable that all the necessary information could be pre-compiled before birth.

Of course, just because we have similarity in results we do not claim that our, rather abstract, robot mechanism is the same or even similar to that which drives the human eye; there can be very many mechanisms that could produce the same behaviour. But what we have shown is that it is possible to learn the necessary information for driving eye saccades very rapidly, and this general principle undermines the nativist assumption that newborn saccading is innate. This means the nativists are obliged to revisit their assumptions and provide better evidence for their stance. The new hypothesis that newborn infants may perform, as yet undetected, rapid learning also offers motivation for biological and psychological investigation to find evidence to either support or refute the idea.

Another aspect concerns the difficulty of observing learning behaviour. Because saccades are such fast movements there is no time for (neural) feedback, and therefore the control must be a forward model [12]. Harris has argued that in this situation “motor babbling” is not just random “neural noise” but is “spontaneous information-gathering action”. We notice that the accurate, learned saccades are often mixed in with more exploratory movements (as gaps in the map are found and filled in). This makes the “motor babbling” much less obvious and it will require careful investigation to discern such learning from behavioural data alone.

Finally, another implication from our model shows the value of proposed mechanisms that may transfer between domains. It is well known that, for saccade errors, undershoot is more common than overshoots. An hypothesis has been proposed [12] that this is an optimum strategy to minimise the total flight-time, because the total flight-time is less with corrective saccades that undershoot as compared with those that overshoot. This effect also occurs in our method: fields near the fovea are predominant among the first to develop because most moves end in such a field, and so when a neighbour is selected it is more likely to be on the near side than the far side of the target. We analysed the data for a run of fixations and found that undershoot occurred in 75% of the cases. Further investigation, with both robotic and biological models may shed some light on the nature of such error patterns and the underlying mechanisms.

6 Conclusion

Facing the problem of induction in empirical sciences we followed Karl Popper's *Logic of Scientific Discovery* where a clear distinction is made between experimental validation and falsification and their role in gaining new and logically valid scientific knowledge. We have outlined how Popper's *Logic of Scientific Discovery* helps us to understand the fundamental problems when robotic experiments are motivated as a test beds for biological models.

The first of our conclusions was that a robot (simulated or real) is not a model. It is an experimental device to test models empirically. Moreover, when asking what kind of models can be tested with robot systems, we had to recognize the loss of the biological domain. The problem of losing the original biological domain means that, under any circumstances, experiments within a robotic domain cannot count as falsification or validation of biological models targeting biological phenomena. Thus, robotics does not directly provide new scientific knowledge for biology. As a consequence robot experiments can only be seen as an inspiration for new biological models and hypotheses. However, we only can accept the robot experiments as biologically relevant if they are explicitly translated back into the biological domain leading to new models and experimental tests. It is exactly this last step which is usually missing when robot experiments are motivated as biologically relevant.

In addition we must admit that there is obviously no formal way for transforming the results from robotics experiments into a biological model and corresponding experimental tests. Thus, robotics will always have a hard time to underline its scientific contribution to biology. In contrast to Barbara Webb we are not interested in tackling this problem by providing guidelines for "good robot models" that lead to "good biological models" [28]. These guidelines can only blur the general demarcation between robotics and biology. This demarcation can hardly be overemphasised since it forces scientists to translate results from robotics experiments back into new biological models and hypothesis amenable for tests in the original biological domain before they can claim that robots experiments are essential part of gaining new scientific knowledge in biology.

In the absence of clearly stated examples of "robotics inspired biology" we have asked for other qualities that provide indirect scientific value for biology. In addition to the obvious quality "Developing new hypotheses and experimental tests", we have stated the following four:

- Biologically inspired implementations
- Proof of concept
- Demonstrations of general principles
- Support of the development of new formalisms, paradigms, and tools

We have seen that any robot implementation of biologically inspired models can only count as a contribution to engineering but not to biology. The very same holds for the so called "existence proofs" or "proofs of concept" in robotics.

The two other qualities serve the development of new formalisms which are believed to provide the more powerful tools for the phenomena not yet understood. However, as long as these tools are not developed yet in such a way that they lead to new hypotheses and experiments they are of course subject of debate and must be carefully motivated [29].

We have discussed four different research activities or programs, namely

- Brain-based devices,
- Conceptual neural evolution,
- Evolving minimal cognitive agents, and
- Developmental robotics.

All these examples have shown that our criticism about the missing "direct biological relevance" is valid. Brain-based devices in particular are motivated as tools that help us to understand how the human brain works. The published results, however, stop at a level where it is claimed that a specific results from the robotic experiments suggest that a certain mechanism "might also be a common property of the neural networks of living animals" [9]. But, this remains an open question no matter how many robotic experiments are performed. Hence, according to our argument, biological relevance of brain-based devices remains open as well.

On the other hand, we have seen that the framework of artificial evolution of minimally cognitive agents is an example for the development of very abstract formalisms without even trying to generate new testable predictions for biological phenomena. As for pure mathematics, no one can tell yet whether this approach will lead to useful tools helping us to gain deeper understanding of biological phenomena. Only time can tell.

It is interesting to notice Arbib's strict view that the models in the framework of "conceptual neural evolution" are not inspired by any approach in robotics, but "will advance the design principles" for intelligent robots [1]. In this sense one could get the impression that Arbib primarily motivates his approach to Brain Research by the current limitations in engineering.

Finally, the extended case study of the learning of a robotic ocular-motor coordination has illustrated the demonstration of two principles. Firstly, the self-calibration of eye saccade movements is necessary and secondly, fast self-calibration is possible. These results gave evidence that saccadic eye behaviour of newborns could be learned rather than innate. Furthermore, with respect to the characteristic saccade errors observed in biological systems a similar dominance of undershooting saccades was found for in our robotic setup. Hence, the dominance of undershoots could be caused by mechanisms present in both systems, human eye and robot. The robotic system allows a formal description of the underlying processes leading to a comprehensive explanation of the characteristic dominance of undershoots. This in turn might help to derive a model explaining this phenomena for biological systems and leading to predictions which are amenable to experimental tests of the biological system. However, as long as this last step (models amenable to experimental test) is missing, we cannot claim biological relevance for our robotics experiments even if the resulting behaviours show similar characteristics.

At this point we conclude this article by stating that computer science or robotics do not belong to empirical sciences because we have shown that in general a robotic experiment can not provide new scientific knowledge about a biological phenomenon. We hope the reader will understand this as good news, since there is no need to limit the design of intelligent information processing systems by specific paradigms or levels of abstraction dictated by biology. The value of robotics and computer science for biology can only become manifested in the demonstration of general principles or the development of new formalisms, paradigms and/or tools. The assumption that biologically motivated robotic implementations directly provide new scientific knowledge for biological phenomena is logically not valid.

ACKNOWLEDGEMENTS

The authors would like to thank Alison Pease and the anonymous reviewers for their valuable and constructive comments. Furthermore, we acknowledge the support by the EU-FP7 projects IM-CLeVeR (ICT-231722) and ROSSI (ICT-216125), and by EPSRC, UK, through grant EP/C516303/1.

REFERENCES

- [1] M. A. Arbib, 'Rana computatrix to human language: towards a computational neuroethology of language evolution', *Phil. Trans. R. Soc. Lond. A*, **361**, 2345–2379, (2003).
- [2] M. A. Arbib and D. H. House, *Depth and detours: an essay on visually guided behavior*, 129–163, Vision, brain and cooperative computation (ed. M. A. Arbib and A. R. Hanson), Cambridge, MA: MIT Press, 1987.
- [3] M. A. Arbib, G. Metta, and P. van der Smagt, *Neurobotics: From Vision to Action*, 1453–1480, Springer Handbook of Robotics (ed. Siciliano and Khatib), Springer, 2008.
- [4] R. Beer, 'The dynamics of active categorical perception in an evolved model agent', *Adaptive Behavior*, **11**(4), 209–243, (2003).
- [5] R. Brooks, 'Intelligence without representation', *Artificial Intelligence*, **47**, 139–159, (1991).
- [6] N. J. Butko, I. R. Fasel, and J. R. Movellan, *Learning about humans during the first 6 minutes of life*, In: Fifth International Conference on Development and Learning. Bloomington, Indiana, USA., 2006.
- [7] F. Chao, M. H. Lee, and J. J. Lee, 'A developmental algorithm for ocular-motor coordination', *Robotics and Autonomous Systems*, DOI information: **10.1016/j.robot.2009.08.002**, (2010).
- [8] H. Chen, K. Matsumoto, J. Ota, and T. Arai, 'Self-calibration of environmental camera for mobile robot navigation', *Robotics and Autonomous Systems*, **55**(3), 177–190, (2007).
- [9] G. M. Edelman, 'Learning in and from brain-based devices', *Science*, **318**(5853), 1103–1105, (2007).
- [10] R. Frigg and S. Hartmann, *Models in Science*, E. N. Zalta (Ed.) Stanford Encyclopedia of Philosophy, 2006.
- [11] L. Hainline, J. Turkel, I. Abramov, E. Lemerise, and C.M. Harris, 'Characteristics of saccades in human infants', *Vision Research*, **24**(12), 1771–1780, (1984).
- [12] C. M. Harris, 'On the optimal control of behaviour: a stochastic perspective', *J Neurosci Methods*, **83**(1), 73–88, (1998).
- [13] C. M. Harris, M. Jacobs, F. Shawkat, and D. Taylor, 'The development of saccadic accuracy in the first seven months', *Clinical vision sciences*, **8**(1), 8596, (1993).
- [14] J. L. Kirchmar and G. M. Edelman, 'Brain-based devices for the study of nervous systems and the development of intelligent machines', *Artificial Life*, **11**, 63–77, (2005).
- [15] E. Kowler and E. Blaser, 'The accuracy and precision of saccades to small and large targets', *Vision Res.*, **35**(12), 1741–1754, (1995).
- [16] D. Muir and J. Field, 'Newborn infants orient to sounds', *Child Development*, **50**(2), 431–436, (1979).
- [17] N. Nagarajan and C. Stevens, 'How does the speed of thought compare for brains and digital computers?', *Current Biology*, **18** (17), 756758, (2008).
- [18] K. Popper, *Logic of Scientific Discovery*, Routledge London, New York, Baltimore, reprint 2009 edn., 2009.
- [19] P. Rochat and T. Striano, 'Emerging self-exploration by 2-month-old infants', *Developmental Science*, **2** (2), 206218, (1999).
- [20] A. Rosenblueth and N. Wiener, 'The role of models in science', *Philos. Sci.*, **12**, 316–321, (1945).
- [21] A. Roucoux, C. Culee, and M. Roucoux, 'Development of fixation and pursuit eye movements in human infants', *Behaviour Brain Research*, **10**, 133–139, (1983).
- [22] B. Russell, *The Problems of Philosophy*, Oxford Paperbacks, 2nd edn., 2001.
- [23] B. Schaal, B. Hopkins, and S. Johnson, 'Prenatal development of post-natal functions', (2005).
- [24] H. Simon, 'Artificial intelligence: An empirical science', *Artificial Intelligence*, **77**, 95–127, (1995).
- [25] J. Triesch, C. Teuscher, G. Deak, and E. Carlson, 'Gaze following: why (not) learn it?', *Developmental Science*, **9** (2), 125147, (2006).
- [26] T. van Gelder, 'The dynamical hypothesis in cognitive science', *BEHAVIORAL AND BRAIN SCIENCES*, **21**(5), 615–628, (1998).
- [27] W. G. Walter, *The Living Brain*, Harmondsworth, Middlesex: Penguin Books, 1961.
- [28] B. Webb, 'Can robots make good models of biological behaviour?', *BEHAVIORAL AND BRAIN SCIENCES*, **24**(6), 1033–1050, (2001).
- [29] B. Webb, 'Animals versus animats: Or why not model the real iguana?', *Adaptive Behavior*, **17**, 269–286, (2009).
- [30] Z. Zhang, 'A flexible new technique for camera calibration', *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, **22** (11), 1330–1334, (2000).